

Investigator Characteristics and Respondent Behavior in Online Surveys ^{*}

Connor Huff[†], Dominika Kruszewska[‡], Christopher Lucas[§],
Anton Strezhnev[¶], Ariel White^{||}

This Draft: October 13, 2015

1 Introduction: A Second Experiment on Researcher Identity in Online Surveys

The experiment discussed in this pre-analysis plan is the second component of a broader project exploring how the perceived identity of researchers affects responses in online surveys. In particular, we analyze how information about the race and gender of the researcher, conveyed in the required consent form and researcher username, affect survey responses. While the theoretical component and hypotheses across these two experiments are identical, this second experiment presents a slightly different way of testing the hypotheses by directly varying the researcher identity via the username used to post the survey.¹ This is in contrast to the first round of the experiment in which we used a generic username called “survey researcher” and only manipulated the researcher name on the consent form.

^{*}Please do not circulate without permission of the authors.

[†]Department of Government, Harvard University, cdezzanihuff@fas.harvard.edu

[‡]Department of Government, Harvard University, dkruszewska@fas.harvard.edu

[§]Department of Government, Harvard University, clucas@fas.harvard.edu

[¶]Department of Government, Harvard University, astrezhnev@fas.harvard.edu

^{||}Department of Government, Harvard University, arwhite@fas.harvard.edu

¹The first experiment, with a more complete discussion of the theory and hypotheses, was also pre-registered on the Experiments in Governance and Politics website here: <http://egap.org/design-registration/registered-designs/>.

The reasons for explicitly manipulating the researcher username are two-fold. First, this is a common practice for researchers conducting online survey experiments. Indeed, several of the authors contributing to this paper explicitly list their name when advertising for their survey. It is useful to know whether information conveyed through this common practice can affect results, in order to establish best practices for how researchers should implement survey experiments on popular online survey platforms. Second, researcher usernames are displayed prominently on online survey platforms such as Amazon’s Mechanical Turk (MTurk). This means that survey respondents are more likely to be cognizant of the name of the researcher when proceeding from the survey posting to the survey itself. If, as we discuss in the related pre-registration document, we expect that perceived researcher identity can affect survey responses, then usernames represent one commonly used means through which respondents can learn the identity of the researcher.

2 Experimental Survey Design

Respondents will be recruited online via Amazon’s Mechanical Turk (MTurk) platform. Many experimental researchers have begun using MTurk for low-cost recruitment of subjects, making it a highly relevant subject pool for investigating the effects of researcher identity within an online context. In online recruitment platforms like Mechanical Turk, accounts have an associated username. We use this username to connote race and gender.

Because the treatment is in the username, rather than in the experiment, subject recruitment must be conducted carefully. In particular, the problem is that the treatment is no longer embedded in the survey (which makes randomization relatively simple) but instead in the request for users to take the survey. We solve this problem as follows.

We pool samples from previous studies conducted by the authors on Mechanical Turk. From this existing pool of respondents, we assign each to one of eight names, listed in Table 1. Subjects are then given a qualification associated with the condition to which they

are assigned. Then, we create separate MTurk accounts for each username and deploy the same survey within each account. Each deployed survey is set such that only MTurk workers with the correct qualification can see the survey (and thus the username associated with it).

To generate the names associated with each of these manipulations, we combined three commonly used lists of racially distinct first and last names. First names were drawn from a combination of lists found in [Bertrand & Mullainathan \(2004\)](#) and [Fryer, Jr. & Levitt \(2004\)](#), while last names were drawn from lists in [Word et al. \(2008\)](#) and [Bertrand & Mullainathan \(2004\)](#). We crossed the lists of first and last names to produce many possible combinations² and drew two names for each of the four manipulation categories (black women, white women, black men, and white men). The full list of names used in this experiment is presented in [Table 1](#).

Black Men	Deshawn Booker	Tyrone Robinson
Black Women	Ebony Gaines	Deja Washington
White Men	Connor Schroeder	Brett Walsh
White Women	Molly Ryan	Laurie Yoder

Table 1: Names used for each of the four investigator name manipulations, based on lists from [Bertrand & Mullainathan \(2004\)](#), [Fryer, Jr. & Levitt \(2004\)](#), [Word et al. \(2008\)](#)

Within the survey, respondents are asked a series of questions about social and political attitudes. We draw questions from Pew, Gallup, and the American National Election Survey specifically asking about issues for which racial and gender cues may prompt different responses. The full text of the outcome questions is presented in [Appendix A](#). We also explore the extent to which respondents are paying attention and are willing to put in effort using attention checks and open text responses. The general structure of the attention checks used in the experiment is taken from [Berinsky et al. \(2014\)](#). Respondents are also asked to complete a randomly assigned writing task, either on their attitudes towards a female president or on a time in their life when they were affected by politics. The latter prompt is sufficiently general that variation in response depth will capture respondent’s general “ef-

²We omitted a few randomly-generated names that already belonged to celebrities, such as Jermaine Jackson.

fort” levels rather than attitudes towards a particular issue. In order to obscure the general purpose of the survey, we randomly permute some of the demographic questions with the outcome questions. However, to avoid priming party affiliation, gender or race, we leave the party ID, gender, and race questions for the end of the survey.

3 Statistical Tests

For Hypothesis 1, we estimate two separate treatment effects.³ The first is the effect of assignment to a putatively female name on the probability that a respondent indicates that they believe that women should have an equal role in the workforce. The second is the effect of assignment to a putatively black name on the respondent’s racial resentment scale. We expect that the effect for the former will be positive while the latter will be negative. For Hypothesis 2, we will estimate the effect of assignment to a putatively white and male name on the probability that a respondent correctly completes both of the attention check assignments. We expect this effect estimate to be positive. For estimation, we will fit a linear probability model of the outcome on treatment and compute standard errors via a nonparametric bootstrapping procedure. While not needed for identification, we will include respondent-level covariates (e.g. gender, income, education) in the regression model in order to increase the efficiency of our estimator.

Our rejection levels for two-sided hypothesis tests of whether the average treatment effects differ from zero are calibrated to correct for problems of multiple testing. We are willing to tolerate an overall Type I error rate of $\alpha = .05$. With three main hypothesis tests, we could obtain a conservative rejection threshold for each individual hypothesis test of $.05/3 = .017$ using the Bonferroni correction. This controls the Familywise Type I Error Rate and guarantees that the probability of any single erroneous rejection in the set of

³Note that the hypotheses referred to in this section reference those laid out in the first preregistration document, which are as follows. H_1 : Assignment to a putatively female/black investigator will increase support for policies that provide for and protect the rights of women/blacks. H_2 : Attention and effort will be greatest among subjects assigned to a putatively white, male investigator.

tests is less than or equal to .05. However, this approach sacrifices a significant amount of power. A less conservative but more powerful approach is to set a rejection threshold to control the False Discovery Rate (FDR). We use the Benjamini-Hochberg procedure to set a rejection level for the hypothesis tests (Benjamini & Hochberg, 1995). This entails a two-step procedure where we order the 3 p-values of the individual hypothesis tests from smallest to largest, $p_{(1)}, \dots, p_{(3)}$ and then set our rejection level to $p_{(k)}$, where k is the largest value of i that satisfies $p_{(i)} \leq \frac{i}{3}\alpha$. This procedure controls the expected share of false hypothesis rejections out of the total number of rejections to be no greater than .05.

We do not specify any ex-ante interactions of the treatment effects with baseline covariates. However, because the mechanism through which any treatment effects operate are of significant interest, we will conduct exploratory analyses of potential treatment effect heterogeneity by estimating models with interactions between treatment and respondent identity variables. Among other interactions, we are interested in seeing whether any average treatment effect is primarily driven by behavior changes among men (in the case of the gender treatment) and white respondents (in the case of the race treatment). We will attempt to replicate any promising results from these exploratory analyses in a follow-up experiment that explicitly registers interactive hypotheses prior to the experiment.

Additionally, because respondents have the option to stop taking the survey after treatment is assigned but before outcomes are measured, there is concern that an analysis conditional on survey completion will be biased for the average treatment effect if treatment also affects the probability that a respondent will drop out. Although it is not possible to adjust for nonignorable drop-out in the absence of prior covariates on respondents, we will examine whether there appears to be systematic differences between treatment arms with respect to attrition and employ sensitivity analyses in the vein of Scharfstein et al. (1999) in order to evaluate the robustness of our estimates to this potential source of bias.

A Appendix: Survey Text

A.1 Informed Consent

Thank you for participating in this survey. Please take time to answer questions honestly and thoroughly. Your responses are essential to our research.

This research is being conducted under the supervision of BLACK/WHITE x MALE/FEMALE NAME at Harvard University. All of the information that we obtain from your session will be anonymous. We do not ask you for your name. Your name or identifying information will not be used in any reports of the research. There will be no direct benefit to you from participation in this study other than the agreed-upon financial compensation. We hope, however, that the research will benefit society by improving our understanding of the factors that influence people's decision making.

If you want to receive the findings of this study, you may contact NAME HERE (GENERIC EMAIL). Complete contact information is as follows.

NAME HERE

Department of Government

Harvard University

Cambridge, MA 02138

GENERIC EMAIL HERE

If you have questions about your rights or about research-related harm, or if your questions, concerns, suggestions, or complaints are not being addressed by the researchers above, please contact:

Director of IRB Operations

Harvard University Committee on the Use of Human Subjects in Research

1414 Massachusetts Avenue, Second Floor

Cambridge, MA 02138

Phone: 617-496-5593

jjaeger@fas.harvard.edu

The nature and purpose of this study have been satisfactorily explained to me and I (participant) agree to become a participant in the study described above. I understand that I am free to discontinue participation at any time if I so choose. Answering yes continues the survey, answering no ends the survey.

Yes

No

A.2 Outcome Questions [Question order randomized]

- Do you agree strongly, agree somewhat, neither agree nor disagree, disagree somewhat, or disagree strongly with these statements?
 - Over the past few years, blacks have gotten less than they deserve.
 - Irish, Italian, Jewish, and many other minorities overcame prejudice and worked their way up. Blacks should do the same without any special favors.
 - It's really a matter of some people not trying hard enough; if blacks would only try harder they could be just as well off as whites.
 - Generations of slavery and discrimination have created conditions that make it difficult for blacks to work their way out of the lower class.

- Some people feel that women should have an equal role with men in running business,

industry and government. Others feel that women's place is in the home. Where would you place yourself on this scale or haven't you thought much about this?

Equal role

Womens place is in the home

Havent thought much about this

- Between now and the 2016 Presidential Election, there will be discussion about the qualifications of presidential candidates - their education, age, race, religion, and so on. If your party nominated a generally well-qualified person for president who happened to be _____, would you vote for that person?

Black ("Yes, would" or "No, would not")

A woman ("Yes, would" or "No, would not")

Catholic ("Yes, would" or "No, would not")

Hispanic ("Yes, would" or "No, would not")

Jewish ("Yes, would" or "No, would not")

Mormon ("Yes, would" or "No, would not")

Gay or lesbian ("Yes, would" or "No, would not")

Muslim ("Yes, would" or "No, would not")

An atheist ("Yes, would" or "No, would not")

- Some people think the government should provide fewer services, even in areas such as health and education, in order to reduce spending. Other people feel that it is important for the government to provide many more services even if it means an increase in spending. Which do you prefer?

Cut services/spending

More services/spending

Demographic Questions [Order permuted with outcome questions]

- What is the highest level of education you have completed?

Less than High School

High School / GED

Some College

2-year College Degree

4-year College Degree

Masters Degree

Doctoral Degree

Professional Degree (JD, MD)

- What is your yearly household income, putting together the income of all the members of your household?

Less than 30,000

30,000 39,999

40,000 49,999

50,000 59,999

60,000 69,999

70,000 79,999

80,000 89,999

90,000 99,999

100,000 or more

Attention Checks [permuted with the outcome questions]

- When a big news story breaks people often go online to get up-to-the-minute details on what is going on. We want to know which websites people trust to get this information. We also want to know if people are paying attention to this survey. To show that you've read this much, please ignore the question and select Reuters website and Huffington Post as your two answers.

When there is a big news story, which is the one news website you would visit first?

(Please only choose one)

New York Times website	The Drudge Report	The Associated Press (AP) website
Huffington Post	Google News	Reuters website
Washington Post website	ABC News website	National Public Radio (NPR) website
CNN.com	CBS News website	USA Today website
FoxNews.com	NBC News website	New York Post Online
MSNBC.com	Yahoo! News	None of these websites

- We are very interested to know what political issues people think are the most relevant today. People often have different attitudes about what issues the United States government should focus on addressing and we would like to understand more about this public debate. We also want to know if people are paying attention to this question. To show that you've read this much, please ignore the question and select Energy and Global trade as your two answers.

Which of the following issues do you think should be the highest priority for President Obama and Congress in 2015? (Please only choose one)

Terrorism	Reducing crime	Energy
Economy	Poor and needy	Influence of lobbyists
Jobs	Military	Transportation
Education	Immigration	Money in politics
Social Security	Environment	Scientific research
Budget deficit	Race relations	Global warming
Health care costs	Moral breakdown	Global trade
Medicare	Tax reform	None of these issues

Open Response Questions

[Randomly assign respondents to one of the two questions below]

- Please write a few sentences about what you think about the United States potentially having a female president.
- Please write a few sentences about a time that politics affected your life.

Demographic Questions [placed at end of survey]

- Generally speaking, do you consider yourself to be a(n):
Democrat
Republican
Independent
- → [if independent] As of today do you lean more to the Republican Party or more to the Democratic Party?
Republican

Democrat

- In talking to people about elections, we often find that a lot of people were not able to vote because they weren't registered, they were sick, or they just didn't have time.

How about you – did you vote in the last presidential election in 2012?

I don't remember.

No, I did not vote.

Yes, I voted.

- → [if yes] Who did you vote for in the last presidential election?

Mitt Romney

Barack Obama

Other (specify)

- What is your ethnicity? Select all that apply.

Black

White

Hispanic

Asian

Native American

Other (specify)

- What is your gender?

Male

Female

Other (please specify)

- What do you think is the purpose of this research?

[Open Response]

- While taking this survey, did you engage in any of the following behaviors?

Use your cell phone

Browse the internet

Talk with another person

Watch TV

Listen to music

Debriefing (sent to all subjects by email AFTER all responses are collected)

The purpose of this study was to learn about how researcher identity influences subjects' responses. In order to learn about this, we needed to manipulate the identity of the researcher. Though the contact information you were given was in fact real, the name of the researcher was not. This study was conducted by Connor Huff, Dominika Kruszewska, Christopher Lucas, Anton Strezhnev, and Ariel White, all at Harvard University. If you have any questions about this study, please contact them at the following location.

Connor Huff, Dominika Kruszewska, Christopher Lucas, Anton Strezhnev, Ariel White
Department of Government
Harvard University
Cambridge, MA 02138

GENERIC EMAIL HERE

References

- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289–300.
- Berinsky, A. J., Margolis, M. F., & Sances, M. W. (2014). Separating the shirkers from the workers? Making sure respondents pay attention on self-administered surveys. *American Journal of Political Science*, 58(3), 739–753. doi: 10.1111/ajps.12081
- Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *American Economic Review*, 94(4), 991–1013. doi: 10.1257/0002828042002561
- Fryer, Jr., R. G., & Levitt, S. J. (2004). The causes and consequences of distinctively Black names. *The Quarterly Journal of Economics*, CXIX(August), 767–806. doi: 10.1162/0033553041502180
- Scharfstein, D. O., Rotnitzky, A., & Robins, J. M. (1999). Adjusting for nonignorable drop-out using semiparametric nonresponse models. *Journal of the American Statistical Association*, 94(448), 1096–1120.
- Word, D. L., Coleman, C. D., Nunziata, R., & Kominski, R. (2008). Demographic aspects of surnames from census 2000. *Unpublished manuscript, Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download>*.